Forum of Mathematics, Sigma (2019), Vol. 7, e46, 12 pages
doi:10.1017/fms.2019.47

A DISTRIBUTION ON TRIPLES WITH MAXIMUM ENTROPY MARGINAL

SERGEY NORIN

Department of Mathematics and Statistics, McGill University, Montréal, Canada;
email: sergey.norin@mcgill.ca

Received 9 January 2017; accepted 17 November 2019

Abstract

We construct an S_3-symmetric probability distribution on $\{(a, b, c) \in \mathbb{Z}^3_{\geqslant 0} : a+b+c = n\}$ such that its marginal achieves the maximum entropy among all probability distributions on $\{0, 1, \ldots, n\}$ with mean $n/3$. Existence of such a distribution verifies a conjecture of Kleinberg *et al.* ['The growth rate of tri-colored sum-free sets', *Discrete Anal.* (2018), Paper No. 12, arXiv:1607.00047v1], which is motivated by the study of sum-free sets.

2010 Mathematics Subject Classification: 11B30 (primary); 05D40 (secondary)

1. Introduction

The recent breakthrough by Croot, Lev and Pach [2] and the subsequent solution of the cap-set problem by Ellenberg and Gijswijt [3] led to a dramatic improvement of known upper bounds on the size of maximum sum-free sets in powers of finite groups. Blasiak et al. [1] extended the Ellenberg–Gijswijt result to multicolored sum-free sets. Even more recently, Kleinberg, Sawin and Speyer [5] established upper bounds for the multicolored version which are essentially tight. Let us state the main result of [5], which motivates our work.

Let p be a prime. A *tri-colored sum-free set in* $\mathbb{F}_p^n$ is a collection of triples $\{(x_i, y_i, z_i)\}_{i=1}^m$ of elements of $\mathbb{F}_p^n$ such that $x_i + y_j + z_k = 0$ if and only if $i = j = k$. Kleinberg, Sawin and Speyer establish an upper bound $m \leqslant e^{\gamma_p n}$ on the size of a tri-colored sum-free set in $\mathbb{F}_p^n$, where γ_p is as follows.

The *entropy* of a probability distribution μ on a finite set I is defined as

$$\eta(\mu) = \sum_{i \in I} \mu(i) \log \mu(i),$$

where we interpret $0 \log 0$ as 0. Let $T = \{(a, b, c) \in \mathbb{Z}^3_{\geqslant 0} : a + b + c = p - 1\}$. Let π be an S_3-symmetric probability distribution on T, and let $\mu(\pi)$ be the marginal probability distribution of π on $[0, p-1]$ corresponding to the first coordinate. (We denote the set of consecutive integers $\{n, n+1, \ldots, n+k\}$ by $[n, n+k]$.) (As π is S_3-symmetric, the choice of a coordinate is irrelevant.) Let γ_p be the maximum entropy of $\mu(\pi)$ over S_3-symmetric probability distributions π on T.

THEOREM 1 (Kleinberg, Sawin and Speyer [5]). *(In the published version of [5], the upper bound part of the statement of Theorem 1, as well as the examples for $n \leqslant 25$ referenced later, were removed, as the proofs of Theorem 2 in an earlier version of this article, and independent work of Pebody [6], showed that they were unnecessary, but they are still available in the referenced arxiv version.) All tri-colored sum-free sets in $\mathbb{F}_p^n$ have size at most $e^{\gamma_p n}$. Moreover, there exist tri-colored sum-free sets in $\mathbb{F}_p^n$ of size at least $e^{\gamma_p n - o(n)}$.*

Every marginal of a symmetric probability distribution on T has mean $(p-1)/3$. Therefore, γ_p is at most the maximum entropy of a probability distribution on $[0, p-1]$ with this mean. The main result of this paper, which was independently obtained by Pebody [6], gives a proof of [5, Theorem 4] showing that the equality holds.

THEOREM 2. *For every $n \geqslant 1$, there exists an S_3-symmetric probability distribution π on $\{(a, b, c) \in \mathbb{Z}^3_{\geqslant 0} : a + b + c = n\}$ such that $\mu(\pi)$ achieves the maximum entropy among probability distributions on $[0, n]$ with mean $n/3$.*

While the definition of γ_p above already implies that it is a computable constant, Theorem 2 provides a much simpler description. As noted in [5], a direct calculation shows that if μ has the maximum entropy among probability distributions on $[0, n]$ with mean $n/3$, then

$$\mu(i) = \frac{\rho^i}{1 + \rho + \cdots + \rho^n}, \tag{1}$$

where ρ is the unique positive real solution to the equation

$$\sum_{i=0}^{n} i\rho^i = \frac{n}{3} \sum_{i=0}^{n} \rho^i. \tag{2}$$

Further, Theorem 2 confirms that the upper bound in Theorem 1 coincides with the bounds established for sum-free sets in [3] and three-colored sum-free sets

https://doi.org/10.1017/fms.2019.47 Published online by Cambridge University Press

in [1]. The value of γ_p is also of interest as it appears in the tight bound for the arithmetic triangle removal lemma of Fox and Lovász [4].

We construct a distribution π satisfying Theorem 2 explicitly. Examples of the distributions satisfying Theorem 2 for $n \leqslant 25$ are provided in [5]. Based on these examples and additional experimentation, we construct a simple S_3-symmetric function of $\{(a, b, c) \in \mathbb{Z}^3_{\geqslant 0} : a + b + c = n\}$ with the marginal given by (1). This construction is presented in Section 2 along with the necessary notation. Unfortunately, the constructed function fails to be nonnegative for $n \geqslant 28$. In Section 3, we modify via a sequence of 'local' changes to establish Theorem 2 in full generality.

2. Notation and the first attempt

Fix positive integer n for the remainder of the paper. Let $T = \{(a, b, c) \in \mathbb{Z}^3_{\geqslant 0} : a+b+c = n\}$. The probabilistic distributions we are interested in form a polytope in $\mathbb{R}^T$, and vectors in $\mathbb{R}^T$ will be the main object of study in the remainder of the paper. We use the convention $\boldsymbol{v} = (v_{abc})_{(a,b,c)\in T}$, that is, we denote by v_{abc} the component of the vector v corresponding to a triple $(a, b, c) \in T$. Let $\{\boldsymbol{e}(a, b, c)\}_{(a,b,c)\in T}$ be the standard basis of $\mathbb{R}^T$. We say that a vector $\boldsymbol{v} \in \mathbb{R}^T$ is *symmetric* if $v_{i_1 i_2 i_3} = v_{i_{\sigma(1)} i_{\sigma(2)} i_{\sigma(3)}}$ for every permutation $\sigma \in S_3$. Let $W \subseteq \mathbb{R}^T$ be the vector space of symmetric vectors. For $(i_1, i_2, i_3) \in T$, define

$$\boldsymbol{s}(i_1, i_2, i_3) = \sum_{\sigma \in S_3} \boldsymbol{e}(i_{\sigma(1)}, i_{\sigma(2)}, i_{\sigma(3)}).$$

The set $\{\boldsymbol{s}(a, b, c) \mid (a, b, c) \in T, a \leqslant b \leqslant c\}$ forms a convenient basis of W. For $\boldsymbol{v} \in \mathbb{R}^T$ and $a \in [0, n]$, define

$$\mu_a(\boldsymbol{v}) = \sum_{i=0}^{n-a} v_{ai(n-a-i)},$$

and let $\mu : \mathbb{R}^T \to \mathbb{R}^{[0,n]}$ be defined by

$$\mu(\boldsymbol{v}) = (\mu_0(\boldsymbol{v}), \mu_1(\boldsymbol{v}), \ldots, \mu_n(\boldsymbol{v})).$$

Note that importantly

$$\mu_i(\boldsymbol{s}(a, b, c)) = 2(\delta_{ia} + \delta_{ib} + \delta_{ic}), \tag{3}$$

for $i \in [0, n]$ and $(a, b, c) \in T$, where δ is the Kronecker delta.

Let ρ de defined by (2). Clearly, $\rho < 1$. Define

$$\boldsymbol{r} = (n, n\rho, \ldots, n\rho^n) \in \mathbb{R}^{[0,n]}.$$

https://doi.org/10.1017/fms.2019.47 Published online by Cambridge University Press

We say that $\boldsymbol{v} \in \mathbb{R}^T$ is *ρ-marginal* if $\mu(\boldsymbol{v}) = \boldsymbol{r}$. Let R denote the set of symmetric, ρ-marginal vectors in $\mathbb{R}^T$. Note that R is an affine space and $R = \boldsymbol{v} + (\mathrm{Ker}(\mu) \cap W)$ for any $\boldsymbol{v} \in R$.

It is easy to see that the next theorem is a reformulation of Theorem 3 using the introduced terminology. (Note that, for convenience, we scaled the target marginal by $n(1 + \rho + \cdots + \rho^n)$.)

THEOREM 3. *There exists a nonnegative vector $\boldsymbol{\pi} \in R$.*

We construct an explicit, albeit not particularly elegant, vector $\boldsymbol{\pi}$ satisfying Theorem 3. As a first step in this section, we construct an auxiliary vector $\boldsymbol{\beta} \in R$, which has a compact description and will be the starting point of the general construction. In Section 3, we finish the construction of $\boldsymbol{\pi}$ by defining a generating set of $\mathrm{Ker}(\mu)$ consisting of vectors with small support and adding an appropriate linear combination of these vectors to $\boldsymbol{\beta}$.

We now define $\boldsymbol{\beta}$. Let

$$\beta_{abc} = \rho^a - \rho^{n-a} + \rho^b - \rho^{n-b} + \rho^c - \rho^{n-c}, \tag{4}$$

for $(a, b, c) \in T$, $a, b, c \geqslant 1$, let

$$\beta_{a0(n-a)} = \sum_{i=1}^{a-1} (-\rho^i + \rho^{n-i}) + \frac{a-1}{2}\rho^a + \frac{n-a+1}{2}\rho^{n-a}, \tag{5}$$

for $1 \leqslant a \leqslant n/2$, let

$$\beta_{00n} = n\rho^n \tag{6}$$

and define the components of $\boldsymbol{\beta}$ for the remaining triples in T by symmetry so that $\boldsymbol{\beta}$ is symmetric.

LEMMA 4. *The vector $\boldsymbol{\beta}$ is ρ-marginal.*

Proof. First, let us note that the identity (5) also holds for $n/2 < a \leqslant n-1$. Indeed, for such i, we have

$$\begin{aligned}
\beta_{a0(n-a)} &= \sum_{i=1}^{n-a-1} (-\rho^i + \rho^{n-i}) + \frac{n-a-1}{2}\rho^{n-a} + \frac{a+1}{2}\rho^a \\
&= \sum_{i=1}^{n-a-1} (-\rho^i + \rho^{n-i}) + \sum_{i=n-a}^{a} (-\rho^i + \rho^{n-i}) + \frac{n-a-1}{2}\rho^{n-a} + \frac{a+1}{2}\rho^a \\
&= \sum_{i=1}^{a-1} (-\rho^i + \rho^{n-i}) - \rho^a + \rho^{n-a} + \frac{n-a-1}{2}\rho^{n-a} + \frac{a+1}{2}\rho^a
\end{aligned}$$

$$= \sum_{i=1}^{a-1}(-\rho^i + \rho^{n-i}) + \frac{a-1}{2}\rho^a + \frac{n-a+1}{2}\rho^{n-a},$$

as desired.

Now we are ready to verify that $\mu_a(\boldsymbol{\beta}) = n\rho^a$ for every $a \in [0, n]$. We have $\mu_n(\boldsymbol{\beta}) = n\rho^n$ by (6). For $1 \leqslant a \leqslant n-1$, we have

$$\begin{aligned}
\mu_a(\boldsymbol{\beta}) &= \sum_{i=0}^{n-a}\beta_{ai(n-a-i)} = \sum_{i=1}^{n-a-1}(\rho^a - \rho^{n-a} + \rho^i - \rho^{n-i} + \rho^{n-a-i} - \rho^{a+i}) \\
&\quad + 2\sum_{i=1}^{a-1}(-\rho^i + \rho^{n-i}) + (a-1)\rho^a + (n-a+1)\rho^{n-a} \\
&= (n-a-1)(\rho^a - \rho^{n-a}) + 2\sum_{i=1}^{n-a-1}(\rho^i - \rho^{n-i}) - 2\sum_{i=1}^{a-1}(\rho^i - \rho^{n-i}) \\
&\quad + (a-1)\rho^a + (n-a+1)\rho^{n-a} \\
&= (n-2)\rho^a + 2\rho^{n-a} + 2(\rho^a - \rho^{n-a}) = n\rho^a,
\end{aligned}$$

as desired. Finally, we have

$$\begin{aligned}
\mu_0(\boldsymbol{\beta}) &= \sum_{i=0}^{n}\beta_{0i(n-i)} \\
&= 2n\rho^n + \sum_{i=1}^{n-1}\left(\sum_{j=1}^{i-1}(-\rho^j + \rho^{n-j}) + \frac{i-1}{2}\rho^i + \frac{n-i+1}{2}\rho^{n-i}\right) \\
&= 2n\rho^n + \sum_{i=1}^{n-1}(-(n-i-1) + (i-1) + i)\rho^i \\
&= \sum_{i=1}^{n}(3i-n)\rho^i = n + \sum_{i=0}^{n}(3i-n)\rho^i = n,
\end{aligned}$$

where the last equality uses (2). □

Upon cursory examination, $\boldsymbol{\beta}$ appears to be a promising candidate for a nonnegative vector in R. In fact, $\boldsymbol{\beta}$ is the only vector in R for $n \leqslant 5$. In general, it is easy to see that $\beta_{abc} \geqslant 0$ for $a, b, c \geqslant 1$. Unfortunately, $\boldsymbol{\beta}$ is nonnegative only for $n \leqslant 27$, whereas $\beta_{0\lfloor n/2\rfloor\lceil n/2\rceil} < 0$ for $n \geqslant 28$. Thus, we have to modify $\boldsymbol{\beta}$ by adding to it an appropriate vector in $\mathrm{Ker}(\mu)$. This is the goal of the next, somewhat technical section.

https://doi.org/10.1017/fms.2019.47 Published online by Cambridge University Press

3. Flattening $\boldsymbol{\beta}$

Given $(a, b, c) \in T$ and $x, y \in \mathbb{N}$ such that $b \geqslant x, c \geqslant x + y$, we define a vector

$$\begin{aligned}\boldsymbol{m}^{x,y}(a, b, c) = &-\boldsymbol{s}(a, b, c) + \boldsymbol{s}(a + x, b - x, c)\\ &- \boldsymbol{s}(a + x + y, b - x, c - y) + \boldsymbol{s}(a + x + y, b, c - x - y)\\ &- \boldsymbol{s}(a + x, b + y, c - x - y) + \boldsymbol{s}(a, b + y, c - y).\end{aligned}$$

CLAIM 5. We have $\boldsymbol{m}^{x,y}(a, b, c) \in \mathrm{Ker}(\mu) \cap W$ for all $(a, b, c) \in T$ and $x, y \in \mathbb{N}$ such that $b \geqslant x, c \geqslant x + y$.

Proof. Clearly, $\boldsymbol{m}^{x,y}(a, b, c) \in W$. Therefore, we only need to show that $\mu_i(\boldsymbol{m}^{x,y}(a, b, c)) = 0$ for every $i \in [0, n]$. This follows immediately from (3), as each Kronecker deltas δ_{ij} for $j \in \{a, a + x, a + x + y, b, b - x, b + y, c, c - y, c - x - y\}$ will appear the same number of times with positive and negative signs in the expansion of $\mu_i(\boldsymbol{m}^{x,y}(a, b, c))$ using (3). □

We think of vectors $\boldsymbol{m}^{x,y}(a, b, c)$ as directions, along which a vector in R can be shifted to obtain another vector in R differing from the original only in a few coordinates. In the remainder of the section, we describe a collection of such shifts which transform $\boldsymbol{\beta}$ into a nonnegative vector.

We will only use a subset of the vectors defined above of the following form. $\boldsymbol{m}(b \to a) = \boldsymbol{m}^{\lceil b/2\rceil, a-b}(0, b, n - b)$ for $2 \leqslant b < a \leqslant n/2$. That is,

$$\begin{aligned}\boldsymbol{m}(b \to a) = &-\boldsymbol{s}(0, b, n - b) + \boldsymbol{s}(\lceil b/2\rceil, \lfloor b/2\rfloor, n - b)\\ &- \boldsymbol{s}(a - \lfloor b/2\rfloor, \lfloor b/2\rfloor, n - a) + \boldsymbol{s}(a - \lfloor b/2\rfloor, b, n - a - \lceil b/2\rceil)\\ &- \boldsymbol{s}(\lceil b/2\rceil, a, n - a - \lceil b/2\rceil) + \boldsymbol{s}(0, a, n - a).\end{aligned} \tag{7}$$

We obtain $\boldsymbol{\pi}$ from $\boldsymbol{\beta}$ by adding to it a linear combination of vectors $\boldsymbol{m}(b \to a)$. The coefficients of these vectors, which we define next, are chosen so that, in particular, $\pi_{0a(n-a)} = \pi_{0b(n-b)}$ for $2 \leqslant a, b \leqslant n - 2$. (Hence, we think of the construction of $\boldsymbol{\pi}$ as 'flattening $\boldsymbol{\beta}$'.) Denote $\beta_{0i(n-i)}$ by z_i for brevity, and let

$$\begin{aligned}c_b &= \frac{2}{n + 1 - 2b}\left(z_b - \frac{1}{n - 1 - 2b}\sum_{i=b+1}^{n-b-1} z_i\right)\\ &= \frac{\sum_{i=b}^{n-b} z_i}{n + 1 - 2b} - \frac{\sum_{i=b+1}^{n-b-1} z_i}{n - 1 - 2b}.\end{aligned}$$

Let $c_{ba} = c_b$ for $2 \leqslant b < a < n/2$ and let $c_{b(n/2)} = \frac{c_b}{2}$ for even n.

https://doi.org/10.1017/fms.2019.47 Published online by Cambridge University Press

We are now ready to define $\boldsymbol{\pi}$:

$$\boxed{\boldsymbol{\pi} = \boldsymbol{\beta} + \sum_{2\leqslant b\leqslant n/2-1} \sum_{b+1\leqslant a\leqslant n/2} c_{ba}\boldsymbol{m}(b \to a)} \tag{8}$$

In the remainder of the section, we show that $\boldsymbol{\pi}$ satisfies Theorem 3. It follows from Claim 5 that $\boldsymbol{\pi} \in R$. Thus, it remains to show that $\pi_{abc} \geqslant 0$ for $(a, b, c) \in T$. This is accomplished in the following series of claims.

CLAIM 6. For $2 \leqslant a \leqslant n/2$, we have

$$\pi_{0a(n-a)} = \frac{n}{n-2}(1 - \rho^n - \rho^{n-1}/2) \geqslant 0. \tag{9}$$

Proof. We have $\pi_{00n} = n\rho^n$, $\pi_{01n} = n\rho^{n-1}/2$ and $\mu_0(\boldsymbol{\pi}) = n$, as $\boldsymbol{\pi}$ is ρ-marginal. Thus, to establish (9), it suffices to verify that $\pi_{0an-a} = \pi_{0bn-b}$ for $2 \leqslant a, b \leqslant n-2$. Define

$$\boldsymbol{\pi}^i = \boldsymbol{\beta} + \sum_{i\leqslant b\leqslant n/2-1} \sum_{b+1\leqslant a\leqslant n/2} c_{ba}\boldsymbol{m}(b \to a).$$

We will show by induction on $n/2 - b$ that

$$\pi^b_{0a(n-a)} = \frac{\sum_{i=b}^{n-b} z_i}{n+1-2b} \tag{10}$$

for every $b \leqslant a \leqslant n - b$. The identity (10) for $b = 2$ implies the claim.

The base case $b = \lfloor n/2 \rfloor$ is trivial. For the induction step and $b < a < n - b$

$$\begin{aligned}
\pi^b_{0a(n-a)} &= \pi^{b+1}_{0a(n-a)} + c_b \\
&= \frac{\sum_{i=b}^{n-b} z_i}{n-1-2b} + \left(\frac{\sum_{i=b}^{n-b} z_i}{n+1-2b} - \frac{\sum_{i=b+1}^{n-b-1} z_i}{n-1-2b}\right) \\
&= \frac{\sum_{i=b}^{n-b} z_i}{n+1-2b},
\end{aligned}$$

as desired, as $z_b = z_{n-b}$. Finally,

$$\begin{aligned}
\pi^b_{0b(n-b)} &= \pi^{b+1}_{0b(n-b)} - \frac{n-1-2b}{2}c_b \\
&= z_b - \frac{n-1-2b}{n+1-2b}\left(z_b - \frac{1}{n-1-2b}\sum_{i=b+1}^{n-b-1} z_i\right) \\
&= \frac{\sum_{i=b}^{n-b} z_i}{n+1-2b},
\end{aligned}$$

finishing the proof of the identity in (10).

https://doi.org/10.1017/fms.2019.47 Published online by Cambridge University Press

It remains to show that $\rho^n + \rho^{n-1}/2 \leqslant 1$. Using (2), we have

$$\frac{n(n+1)}{3}\left(\rho^n + \frac{\rho^{n-1}}{2}\right) \leqslant n\rho^n + \frac{n(n-1)}{2}\rho^{n-1}$$
$$\leqslant \sum_{i=0}^{n} i\rho^i = \frac{n}{3}\sum_{i=0}^{n}\rho^i \leqslant \frac{n(n+1)}{3},$$

implying the desired inequality. □

We now proceed to establish the estimates which will allow us to prove the nonnegativity of $\boldsymbol{\pi}$. We start with a couple of indirect bounds on ρ.

CLAIM 7. We have

$$\rho^{n/2+1} \geqslant \frac{2}{3e}. \tag{11}$$

Proof. It can be verified by a computer that $\rho^{n/2+1} \geqslant 1/3$ for $n \leqslant 15$. Thus, (11) holds for $n \leqslant 15$, and we assume $n \geqslant 16$.

Let $\alpha = (n+2)(1-\rho)/\rho$. Then $\rho = (n+2)/(n+2+\alpha)$, and

$$\rho^{n+1} \geqslant \rho^{n+2} = \frac{1}{(1+\alpha/(n+2))^{n+2}} \geqslant e^{-\alpha}. \tag{12}$$

We claim that $\alpha \leqslant 2\log(3e/2)$. Note that by (12), this claim implies (11).

We start the proof of the claim by multiplying both sides of (2) by $(1-\rho)^2$, expanding and rearranging terms to obtain

$$(n+3)\rho - n - \rho^{n+1}((2n+3) - 2n\rho) = 0.$$

Using (12) to bound ρ^{n+1} and otherwise expressing ρ in terms of α, we get

$$\frac{(n+3)(n+2)}{n+2+\alpha} - n - e^{-\alpha}\left(2n+3 - \frac{2n(n+2)}{n+2+\alpha}\right) \geqslant 0.$$

Multiplying the above inequality by $n+2+\alpha$, we obtain

$$3n + 6 - n\alpha - e^{-\alpha}(3n + 6 + (2n+3)\alpha) \geqslant 0.$$

Let $f(x,n) = 3n + 6 - nx - e^{-x}(3n + 6 + (2n+3)x)$. We have $f(2\log(3e/2), 16) = -0.14055\ldots$, and $\frac{\partial f}{\partial x}(x,n)$ and $\frac{\partial f}{\partial n}(x,n)$ are easily seen to be negative for $x \geqslant 2\log(3e/2)$ and $n \geqslant 16$. Thus, $f(x,n) < 0$ for all $x \geqslant 2\log(3e/2)$ and $n \geqslant 16$. As $f(\alpha, n) \geqslant 0$, we have $\alpha < 2\log(3e/2)$, as desired. □

https://doi.org/10.1017/fms.2019.47 Published online by Cambridge University Press

Define

$$\delta = \frac{1-\rho}{e\rho}.$$

CLAIM 8. We have

$$(i+1)(1-\rho)^2\rho^i \leqslant \delta,$$

for all $i \geqslant 0$.

Proof. By the arithmetic mean–geometric mean inequality, we have

$$\left(\frac{\rho}{i+1}\right)^{i+1}(1-\rho) \leqslant \left(\frac{1}{i+2}\right)^{i+2}.$$

Therefore,

$$(i+1)(1-\rho)^2\rho^i \leqslant \frac{(1-\rho)}{\rho}\left(\frac{i+1}{i+2}\right)^{i+2} \leqslant \frac{(1-\rho)}{\rho}\frac{1}{e} = \delta,$$

as desired. □

Next we estimate c_b. Define $\Delta_b = z_b - z_{b+1}$. Direct calculation shows that

$$\Delta_b = \tfrac{1}{2}((b+1)\rho^b - b\rho^{b+1} + (n-b-1)\rho^{n-b} - (n-b)\rho^{n-b-1}).$$

CLAIM 9. We have

$$\Delta_{i+1} \leqslant \Delta_i \leqslant \Delta_{i+1} + \delta$$

for all $1 \leqslant i \leqslant n-1$.

Proof. We have

$$\begin{aligned} 2(\Delta_i - \Delta_{i+1}) &= (i+1)\rho^i - (2i+2)\rho^{i+1} + (i+1)\rho^{i+2} \\ &\quad + (n-i-1)\rho^{n-i-2} - 2(n-i-1)\rho^{n-i-1} + (n-i-1)\rho^{n-i} \\ &= (1-\rho)^2((i+1)\rho^i + (n-i-1)\rho^{n-i-2}). \end{aligned}$$

The last term is clearly nonnegative, and it is at most 2δ by Claim 8. Thus, the claim holds. □

CLAIM 10. We have

$$0 \leqslant \Delta_i \leqslant \frac{\delta(n-1-2i)}{2},$$

for all integers $1 \leqslant i \leqslant (n-1)/2$.

Proof. As $z_i = z_{n-i}$, we have $\Delta_i = -\Delta_{n-1-i}$ for all i. Thus, if n is odd, we have $\Delta_{(n-1)/2} = 0$. For even n, we have $\Delta_{n/2} = -\Delta_{n/2-1}$, and $\Delta_{n/2} \leqslant \Delta_{n/2-1} \leqslant \Delta_{n/2}+\delta$

https://doi.org/10.1017/fms.2019.47 Published online by Cambridge University Press

by Claim 9. Thus, $0 \leqslant \Delta_{n/2-1} \leqslant \delta/2$. This establishes the claim for $i \in \{(n-1)/2, n/2-1\}$.

The claim for general i follows directly from Claim 9 by induction on $\lfloor n/2 \rfloor - i$, with the result of the preceding paragraph used as the base case. □

CLAIM 11. We have

$$c_b \leqslant \frac{\delta(n-2b)}{6} \tag{13}$$

for every positive integer $2 \leqslant b \leqslant n/2-1$.

Proof. As in Claim 10, the proof is by induction on $\lfloor n/2 \rfloor - b$, and the base case is $b = \lfloor n/2 \rfloor - 1$.

Suppose first that n is even. Then $n = 2b+2$ in the base case, and $c_b = 2\Delta_b/3$ by definition. As $0 \leqslant \Delta_b \leqslant \delta/2$ by Claim 10, (13) holds. If n is odd, then $n = 2b+3$, $c_b = \Delta_b/2$ and $0 \leqslant \Delta_b \leqslant \delta$ by Claim 10, implying that (13) once again holds. This finishes the proof of the base case.

For the induction step, note that

$$\begin{aligned}
&\frac{(n+1-2b)(n-1-2b)c_b}{2} \\
&\quad = (n-1-2b)z_b - -\sum_{i=b+1}^{n-b-1} z_i \\
&\quad = (n-1-2b)\Delta_b + (n-1-2b)z_{b+1} - \sum_{i=b+1}^{n-b-1} z_i \\
&\quad = (n-1-2b)\Delta_b + (n-1-2(b+1))z_{b+1} - \sum_{i=b+2}^{n-b-2} z_i \\
&\quad = (n-1-2b)\Delta_b + \frac{(n-1-2b)(n-3-2b)c_{b+1}}{2}.
\end{aligned}$$

Thus,

$$c_b = \frac{2\Delta_b + (n-3-2b)c_{b+1}}{(n+1-2b)}.$$

Using the bounds on Δ_b from Claim 10 and the induction hypothesis applied to c_{b+1}, we obtain

$$0 \leqslant c_b \leqslant \frac{\delta(n-1-2b) + \delta(n-3-2b)(n-2-2b)/6}{n+1-2b} = \frac{\delta(n-2b)}{6},$$

as desired. □

CLAIM 12. We have

$$\rho^b - \rho^{n-b} \geqslant 4c_b$$

for all positive integers $2 \leqslant b \leqslant n/2 - 1$.

Proof. Let

$$f(x) = \rho^x - \rho^{n-x} - \frac{2\delta(n-2x)}{3}.$$

By Claim 11, it suffices to show that $f(x) \geqslant 0$ for all $x \leqslant n/2$. As $f(n/2) = 0$, and $f''(x) \geqslant 0$ for $x \leqslant n/2$, it is enough to verify that $f'(n/2) \leqslant 0$, that is,

$$-2\rho^{n/2}\log\rho \geqslant \frac{4}{3}\delta = \frac{4(1-\rho)}{3e\rho}.$$

As $-\log\rho \geqslant 1 - \rho$, the above is implied by Claim 7. □

The next claim finishes the proof of Theorem 3.

CLAIM 13. We have $\pi_{xyz} \geqslant 0$ for $(x, y, z) \in T$, $x, y, z \geqslant 1$.

Proof. Assume that $x \leqslant y \leqslant z$. Suppose that the component corresponding to (x, y, z) is negative in $\boldsymbol{m}(b \to a)$ for some $2 \leqslant b < a \leqslant n/2$. Direct examination of (7) shows that, if $z > n/2$, then

(N1) either $b \in \{2x, 2x+1\}$ and $a = x + y$, in which case $(x, y, z) = (\lfloor b/2\rfloor, a - \lfloor b/2\rfloor, n - a)$, or

(N2) $b \in \{2x, 2x-1\}$, $a = y$, in which case $(x, y, z) = (\lceil b/2\rceil, a, n-a-\lceil b/2\rceil)$.

If $z < n/2$, then

(N2) either $b \in \{2x, 2x-1\}$ and $a = y$, $(x, y, z) = (\lceil b/2\rceil, a, n-a-\lceil b/2\rceil)$ as before or

(N3) $b \in \{2x, 2x-1\}$, $a = z$, in which case $(x, y, z) = (\lceil b/2\rceil, n-a-\lceil b/2\rceil, a)$.

If $z = n/2$, then any of the above cases can potentially occur, but in cases (N1) and (N3), the component of $\boldsymbol{m}(b \to a)$ corresponding to (x, y, z) is equal to $c_b/2$ rather than c_b.

Suppose first that $x < y < z$. By the above analysis, the total negative contribution to π_{xyz} of vectors $\boldsymbol{m}(b \to a)$ for $2 \leqslant b < a \leqslant n/2$ is at most $4\max\{c_{2x}, c_{2x+1}, c_{2x-1}\} \leqslant \rho^x - \rho^{n-x}$, where the last inequality holds by Claim 12. Thus,

$$\pi_{xyz} \geqslant \beta_{xyz} - \rho^x - \rho^{n-x} \geqslant (\rho^z - \rho^{n-y}) + (\rho^y - \rho^{n-z}) > 0,$$

as desired.

https://doi.org/10.1017/fms.2019.47 Published online by Cambridge University Press

Finally, suppose that $x \leqslant y \leqslant z$, and one of these inequalities is nonstrict. Then a vector $\boldsymbol{m}(b \to a)$ can only contribute negatively to π_{xyz} if $x < y = z$ and $(x, y, z) = (\lceil b/2 \rceil, a, n-a-\lceil b/2 \rceil)$. Note that in this case, the component of $\boldsymbol{m}(b \to a)$ corresponding to (x, y, z) is equal to $2c_b$, rather than c_b, but the bound $4\max\{c_{2x}, c_{2x+1}, c_{2x-1}\}$ on the total negative contribution established in the previous case still holds. □

Acknowledgements

We would like to thank the anonymous referee and Lisa Sauermann for pointing out an error in the estimates used in the proof of Theorem 3 in an earlier version of this paper. The author is supported by an NSERC Discovery grant.

Conflict of Interest: None

References

[1] J. Blasiak, T. Church, H. Cohn, J. A. Grochow, E. Naslund, W. F. Sawin and C. Umans, ‘On cap sets and the group-theoretic approach to matrix multiplication’, *Discrete Anal.* (2017), Paper No. 3, 27 pp.

[2] E. Croot, V. F. Lev and P. P. Pach, ‘Progression-free sets in $\mathbb{Z}_4^n$ are exponentially small’, *Ann. of Math. (2)* **185**(1) (2017), 331–337.

[3] J. S. Ellenberg and D. Gijswijt, ‘On large subsets of $\mathbb{F}_q^n$ with no three-term arithmetic progression’, *Ann. of Math. (2)* **185**(1) (2017), 339–343.

[4] J. Fox and L. M. Lovász, ‘A tight bound for Green’s arithmetic triangle removal lemma in vector spaces’, *Adv. Math.* **321** (2017), 287–297.

[5] R. Kleinberg, W. Sawin and D. E. Speyer, ‘The growth rate of tri-colored sum-free sets’, *Discrete Anal.* (2018), Paper No. 12, arXiv:1607.00047v1, 10 pp.

[6] L. Pebody, ‘Proof of a conjecture of Kleinberg–Sawin–Speyer’, *Discrete Anal.* (2018), Paper No. 13, 7 pp.